Convergence and stability of discretization methods for functional equations

By J. M. Watt*

This paper gives an expository account of recent advances in the study of discretization methods for solving functional equations. Examples of the application of the theory to both initial value, and two-point boundary value problems for differential equations are included.

(First received October 1967)

Interesting advances have been made recently in the theoretical study of discretization methods for the solution of functional equations. These have achieved great generality by using the notation of functional analysis to achieve an economy of exposition similar to that given by the use of matrices.

This paper, which was originally given in a slightly different form at the symposium on the 'Numerical Solution of Differential Equations' at St. Andrews in June 1967, aims to give a connected exposition of this subject and to indicate its use by applying the results to particular examples of methods for solving differential equations.

1. Normed linear spaces: Banach spaces

In this paper I assume that the notion of a linear (or vector) space is familiar to the reader.

In dealing with a single real or complex quantity one can take its modulus $|x|$ as a measure of its magnitude. It is also useful to define measures of magnitude of elements of more general linear spaces. If they satisfy suitable conditions we call these measures of magnitude *norms* and denote them by $||x||$.

In order for a norm to satisfy our usual ideas for a measure of magnitude we require:

(1) $||x|| \geqslant 0$. $||x|| = 0$ if and only if $x = 0$.

(2) $||\alpha x|| = |\alpha| . ||x||$ $\quad \alpha$ a real number.

(3) $||x + y|| \leqslant ||x|| + ||y||$ Triangle inequality.

As norm in a finite dimensional vector space we can take

(*a*) $||\boldsymbol{v}||_\infty = \max |v_i| \qquad 1 \leqslant i \leqslant n$

This is the Chebyshev, L_∞ or uniform norm.

(*b*) $||\boldsymbol{v}||_2 = (|v_1|^2 + \ldots + |v_n|^2)^{1/2}$

The L_2 or Euclidean norm.

(*c*) $||\boldsymbol{v}||_1 = |v_1| + \ldots + |v_n|$

The L_1 norm.

Corresponding to these we have as norms on the linear space $C[a, b]$ of all continuous functions on $[a, b]$

(*a*) $||f||_\infty = \max |f(x)| \qquad a \leqslant x \leqslant b$

(*b*) $||f||_2 = \left(\int_a^b |f(x)|^2 dx \right)^{1/2}$

(*c*) $||f||_1 = \int_a^b |f(x)| \, dx.$

Convergence of sequences is defined using the Cauchy criterion. That is (u_n) converges if given $\epsilon > 0$ we can find N such that $||u_n - u_m|| < \epsilon$ provided $n \geqslant N, m \geqslant N$.

If each convergent sequence has a limit in the space, the space is said to be complete. A complete normed linear space is called a Banach space.

If norms are defined on two Banach spaces A and B, and L is a continuous linear mapping (or linear operator or function) from A to B, that is $L(\alpha x + \beta y) = \alpha Lx + \beta Ly$, we can define a norm of L as

$$||L|| = \sup ||Lx||_B \quad \text{for all } x \text{ such that } ||x||_A = 1.$$

For more details the books of Simmons (1963), Dieudonne (1960), Liusternik and Sobolev (1961) and Collatz (1966) should be consulted (in this order). Simmons, however, does not treat differentiation; the best account is in Dieudonne.

2. Differentiation of mappings in Banach space

If $f(x)$ is a real function with real argument, we have

$$f(x + h) = f(x) + hf'(x) + O(h^2).$$

We use this equation as the *definition* of the derivative of a function with respect to a vector or function argument. Higher derivatives are similarly defined as the coefficients in Taylor series.

Examples

(1) Vector argument.

$$f(x_1 + h_1, \ldots, x_n + h_n) = f(x_1, \ldots, x_n) + \left[\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right][h_1, \ldots, h_n]^T + O(||h||)^2$$

So the derivative is the vector $\left[\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right]$.

* *Department of Computational and Statistical Science, The University, Liverpool* 3.

(2) Vector function, vector argument.

$f_i(x_1, \ldots, x_n)$ for $i = 1, \ldots, n$.

$$f(x+h) = f(x) + \begin{bmatrix} \frac{\partial f_1}{\partial x_1}, \ldots, \frac{\partial f_1}{\partial x_n} \\ \text{--------} \\ \frac{\partial f_n}{\partial x_1}, \ldots, \frac{\partial f_n}{\partial x_n} \end{bmatrix} \begin{bmatrix} h_1 \\ -- \\ h_n \end{bmatrix} + O(||h||^2)$$

$$= f(x) + Df(x)h + O(||h||^2)$$

So the derivative is the matrix of partial derivatives of the functions.

(3) Function argument.

Consider the mapping F taking the function $y(x)$ into another function $F(y(x))$ defined by

$$F(y(x)) = \frac{dy(x)}{dx} - f(y(x)).$$

Now

$$\begin{aligned} F(y(x) + h(x)) &= y'(x) + h'(x) - f(y(x) + h(x)) \\ &= y'(x) + h'(x) - f(y(x)) \\ &\quad - f_y(y(x)).h(x) + O(||h||^2) \end{aligned}$$

so that

$$DF(y(x))h(x) = h'(x) - f_y(y(x)).h(x)$$

and

$$DF(y(x)) = \frac{d}{dx} - f_y(y(x)).$$

(4) Second derivative of vector function with vector argument.

$$\begin{aligned} f_i(x+h) &= f_i(x) + \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(x)h_j \\ &\quad + \frac{1}{2!}\sum_{j,k=1}^{n} \frac{\partial^2 f_i(x)}{\partial x_j \partial x_k} \cdot h_j h_k + O(||h||)^3 \end{aligned}$$

so

$$D^2 f(x)h^2 = \left\{\frac{1}{2!}\sum_{j,k=1}^{n} \frac{\partial^2 f_i(x)}{\partial x_j \partial x_k} h_j h_k\right\}.$$

It is normally the differential, and not the derivative that is needed, and this is often easier to define (cf. $DF(y(x))$ above).

Note that the first derivative is a linear function of h and the second is a symmetric bilinear function, etc.

As another example consider the perturbation of the solution of a differential equation. Consider a mapping F from a space of continuously differentiable functions on $[a, b]$ to the product of the space of reals and the space of continuous functions on $[a, b]$. (Or in a short-hand notation G: $C'[a, b] \to R \times C[a, b]$.)

$$G(y) = \begin{cases} y(a) - c \\ y'(x) - f(y(x)) & x \in [a, b] \end{cases}$$

$G(y) = 0$ has a unique solution $y(x)$ if f satisfies a Lipschitz condition.

Define $z(x)$ by

$$G(z) = \alpha\Phi \text{ where } \Phi = \begin{cases} \gamma \\ \phi(x) \end{cases}.$$

Putting $z(x) = y(x) + \alpha e(x) + O(\alpha^2)$ we get

$$\alpha\Phi = G(z) = G(y) + DG(y)\alpha e + O(\alpha^2)$$

So we deduce that

$$DG(y).e = \Phi$$

or

$$e(a) = \gamma$$

$$e'(x) - f_y(y(x))e(x) = \phi(x).$$

This determines to the first order in α the perturbation in $z(x)$ due to a perturbation $\alpha\Phi$ in the differential equation.

It is trivial to extend this theory to that of a differential equation in a Banach space E. We merely take the values of $y(x)$, c and $f(y(x))$ to be elements of E, so that G is a mapping from the differentiable subspace of the Banach space $C_E[a, b]$ of continuous functions on $[a, b]$ with values in E (with norm $||z|| = \sup_{a \leqslant x \leqslant b} ||z(x)||_E$), onto the product space $E \times C_E[a, b]$ (with norm $||w|| = \max(||w_0||_E, \sup_{a \leqslant x \leqslant b} ||w_1(x)||_E)$ if $w = (w_0, w_1)$).

If E is taken to be the vector space of m dimensions with real coordinates the theory applies immediately to systems of m differential equations in m unknown real functions.

Notice that the estimates of the discretization error for the initial value problem given by Henrici (1962) are of the above form.

Notation: Arguments of functions (or operators) are usually enclosed in parentheses, but if the function is linear the parentheses are often omitted. Linear functions, and the multiplication of two functions are distinguished if necessary by the context.

If an equation holds for all allowed values of an argument, it is usual to omit the argument. Thus the equation for $e(x)$ above would be written

$$e' - f_y(y)e = \phi.$$

3. Discretizations

In solving problems numerically we can deal only with finite dimensional spaces. Infinite dimensional problems such as the initial value problem just considered must therefore be approximated by a discrete finite dimensional problem. This process has come to be known as *discretization.*

Considering the initial value problem again we define the discretization operator Δ_A^N, which is a linear— mapping, from $C'[a, b]$ to R^{N+1} by

$$\Delta_A^N y = [y(a), y(a+h), y(a+2h), \ldots, y(b)]$$

$$h = h_N = (b-a)/N.$$

Similarly Δ_B^N from $R \times C[a, b]$ to R^{N+1} is defined by

$$\Delta_B^N[z_0, z_1] = [z_0, z_1(a), z_1(a+h), \ldots, z_1(a+(N-1)h)]$$

$$z_0 \in R, z, \in C[a, b].$$

The original equation $G(y) = 0$, or more generally $G(y) = b$ is now replaced by

$$\Phi_N(\eta^N) = \Delta_B^N b.$$

Here $\Delta_B^N b$ and η^N are vectors of $(N + 1)$ components, and this equation is a set of $(N + 1)$ equations in the $(N + 1)$ unknown components of η^N. The relationship between the various spaces is illustrated in **Fig. 1**.

The mapping Φ_N is chosen so that Φ_N is a good approximation to G, or rather so that

$\Phi_N(\Delta_A^N z)$ is a good approximation to $\Delta_B^N G(z)$.

Each of these elements of B^N is a function of an element of A.

We expect the approximation to get better as N increases and tends to infinity.

The discretization is said to be *p-consistent* if there is a number C such that

$$||\Phi_N(\Delta_A^N y) - \Delta_B^N G(y)|| \leqslant C N_0^{-p}. \tag{1}$$

For example—the initial value problem solved by Euler's method:

$$\eta^N = [\eta_0^N, \eta_1^N, \ldots, \eta_N^N]$$

$$0 = \Phi_N(\eta^N) = \begin{cases} \eta_0^N - c \\ (\eta_i^N - \eta_{i-1}^N)/h - f(\eta_{i-1}^N) \end{cases}$$
$$1 \leqslant i \leqslant N$$

We have

$$\max_i |(y(x_i^N) - y(x_{i-1}^N))/h - f(y(x_{i-1}^N)) - [y'(x_{i-1}^N) - f(y(x_{i-1}^N))]|$$

$\leqslant Lh \leqslant L(b - a)/N$ for some L provided $f(y)$ is sufficiently smooth. Hence Euler's method is 1-consistent.

We are interested in the behaviour of $||\eta^N - \Delta_A^N y||$ as $N \to \infty$. The method is convergent if this tends to zero, for then $\eta_i^N \to y(x_i^N)$ as $N \to \infty$. The method is *convergent of order p* is there is a number C_1 such that

$$||\eta^N - \Delta_A^N y|| \leqslant C_1 N^{-p} \quad \text{for all } N. \tag{2}$$

It is, of course, assumed that Φ_N has a unique inverse so that $\Phi_N(\eta^N) = \Delta_B^N b$ can be solved for each value of N. Thus $\eta^N = \phi_N^{-1}(\Delta_B^N b)$. It may happen that this set of equations becomes more and more ill-conditioned as N tends to infinity, and in the limit that it becomes 'infinitely ill-conditioned'. If this is the case we say that the method, or the sequence of mappings (Φ_N^{-1}), is unstable.

On the contrary if there is a number L such that

$$||\Phi_N(\zeta_1^N) - \Phi_N(\zeta_2^N)|| \geqslant L||\zeta_1^N - \zeta_2^N|| \tag{3}$$

for all ζ_1^N, ζ_2^N and for all N we say that (Φ_N^{-1}) is asymptotically stable. (The condition is that Φ_N^{-1} satisfies a uniform Lipschitz condition.) Since only asymptotic stability is used in this paper the adjective 'asymptotic' is usually omitted.

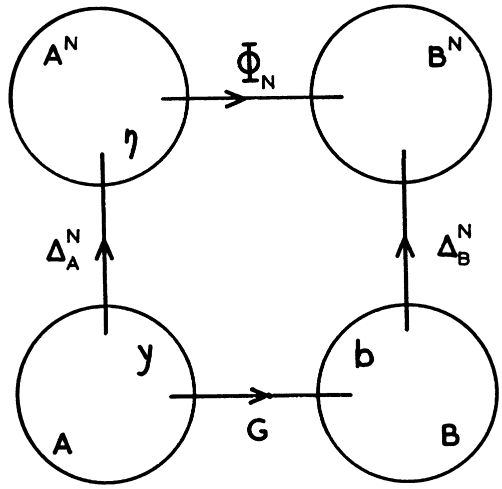

Fig. 1

Remarks

(*a*) According to this definition, stability is a property of the discretized problem only; it does not depend on the original problem.

(*b*) If rounding error is present and instead of $\Phi_N(\eta^N) = \Delta_B^N(b)$ we have $\Phi_N(\eta^N) = \Delta_B^N(b) + \epsilon^N$, the stability guarantees that the effect of the small rounding errors ϵ^N on η^N, is small.

(*c*) Our definition of stability includes both Henrici's (1962) strongly stable and weakly stable methods. (For an interesting discussion of strong and weak stability see Stetter (1965a).)

(*d*) Most stability definitions that have been given for ordinary differential equations are what I would consider to be necessary and sufficient conditions for stability.

It can now be shown that with these definitions, consistency and stability imply convergence.

Theorem If Φ_N is consistent of order p at y and the sequence of inverse mappings (Φ_N^{-1}) is stable, then the discretization Φ_N of G is convergent of order p at $b = Gy$.

Proof

$$||\eta^N - \Delta_A^N y||$$

$$\leqslant \frac{1}{L}||\Phi_N(\eta^N) - \Phi_N(\Delta_A^N y)|| \quad \text{(by stability (3)).}$$

$$= \frac{1}{L}||\Delta_B^N G(y) - \Phi_N(\Delta_A^N y)||$$

$$\text{(since } \Phi_N(\eta^N) = \Delta_B^N b = \Delta_B^N G(y))$$

$$\leqslant \frac{C}{L} N^{-p} \quad \text{(by } p\text{-consistency (1)).}$$

Hence p-convergence is proved.

Stetter (1965) has proved a general theorem about the form of the asymptotic discretization error; his theorem and the one above can be used to derive easily many of the results in Henrici's (1962) book on Discrete Variable Methods.

Theorem (Stetter) $G(y) = b$ is discretized with $\Phi_N(\eta^N) = \Delta_B^N b$ and there is a relation

$$\Phi_N(\Delta_A^N z) = \Delta_B^N \left\{G(z) + \sum_{r=1}^{p} g_r(z)N^{-r}\right\} + O(N^{-p-1}) \quad (4)$$

which holds for all z in a subset A_1 of A containing a neighbourhood of y, and where $g_r(y) = 0$, $r < p$.

The sequence (Φ_N^{-1}) is stable, G and the g_r are differentiable, and $G'(y)e = d$ has a unique solution e in A_1 for all d in B.

Then

$$\eta^N = \Delta_A^N(y + N^{-p}e) + O(N^{-p-1})$$

where e is given by

$$G'(y)e = -g_p(y). \quad (5)$$

PROOF Put $\epsilon^N = \eta^N - \Delta_A^N(y + N^{-p}e)$

Then

$$\begin{aligned} L\|\epsilon^N\| &\leqslant \|\Phi_N(\eta^N) - \Phi_N(\Delta_A^N(y + N^{-p}e))\| \text{ (by stability)} \\ &\leqslant \left\|\Delta_B^N\left\{b - G(y + N^{-p}e) - \sum_{r=1}^{p} g_r(y + N^{-p}e)\right\}\right\| \\ &\qquad + O(N^{-p-1}) \\ &\qquad \text{(by (4) provided } N \geqslant N_0 \text{ so that } y + N^{-p}e \text{ is in } A_1) \\ &= \|\Delta_B^N\{b - G(y) - G'(y)N^{-p}e - g_p(y)N^{-p}\}\| \\ &\qquad + O(N^{-p-1}) \\ &\qquad \text{(using the differentiability)} \\ &= O(N^{-p-1}) \qquad \text{(using (5) and } G(y) = b). \end{aligned}$$

Hence $\|\epsilon^N\| = O(N^{-p-1})$ and the theorem is proved.

The theorem can be extended to obtain further terms in the expansion of η^N provided that G and g_r are sufficiently differentiable. This justifies the extrapolation to the limit techniques which have been treated in several recent papers: see Gragg (1965), Bulirsch and Stoer (1966, 1966a). It also extends to the case where the powers of N in (4) are not integral.

Notice that the theorem derives information about an expansion of the inverse of an operator G from an expansion of G itself. It is useful when G^{-1} is required but G is simpler to deal with.

This theory is concerned with the behaviour of discretizations as N tends infinity. In practice, of course, a finite value of N must be used.

Example 1: *Second order Runge Kutta*

A_1 is the continuously differentiable subspace of the space $A = C_E[a, b]$ of continuous functions on $[a, b]$ with values of E.

B is the space of pairs of elements the first of which belongs to E and the second to $C_E[a, b]$.

$A^N = B^N$ is the space of vectors of $(N + 1)$ components belonging to E.

Δ_A^N: a mapping from A to A^N defined by

$$\Delta_A^N z = \{z(x_0), z(x_1), \ldots, z(x_N)\}$$
$$x_n = a + nh_m, \quad h_N = (b - a)/N$$

Δ_B^N: a mapping from B to B^N defined by

$$\Delta_B^N\{z_0, z_1\} = \{z_0, z_1(x_0), z_1(x_1), \ldots, z_1(x_{N-1})\}$$

$$G(z): A \to B \text{ defined by } G(z) = \begin{cases} z(a) - c \\ z'(x) - f(z(x)) \quad x \in [a, b] \end{cases}$$

$\Phi_N(\zeta^N)$: $A^N \to B^N$

$$\Phi_N(\zeta^N) = \begin{cases} \zeta_0^N - c \\ (\zeta_{n+1}^N - \zeta_n^N)/h_N - \tfrac{1}{2}[f(\zeta_n^N) + f(\zeta_n^N + h_N f(\zeta_n^N))] \\ \qquad 0 \leqslant n < N. \end{cases}$$

We have

$$\begin{aligned} \Phi_N(\Delta_A^N z) &= \begin{cases} z(a) - c \\ (z(x_{n+1}) - z(x_n))/h_N \\ \qquad - \tfrac{1}{2}[f(z(x_n)) + f(z(x_n) + h_N f(z(x_n)))] \end{cases} \\ &= \begin{cases} z(a) - c \\ [z'(x_n) - f(z(x_n))] \\ \qquad + \tfrac{1}{2}h_N[z''(x_n) - f'(z(x_n))f(z(x_n))] \\ \qquad + h_N^2[\tfrac{1}{6}z'''(x_n) - \tfrac{1}{4}f''(z(x_n))(f(z(x_n)))^2] \\ \qquad + O(h_N^3) \end{cases} \\ &= \Delta_B^N\{G(z) + g_1(z)/N + g_2(z)/N^2\} + O(N^{-3}) \end{aligned}$$

where

$$g_1(z) = \begin{cases} 0 \\ \tfrac{1}{2}(b - a)[z'' - f'(z)f(z)] \end{cases}$$

$$g_2(z) = \begin{cases} 0 \\ (b - a)^2[\tfrac{1}{6}z''' - \tfrac{1}{4}f''(z)(f(z))^2]. \end{cases}$$

If $G(y) = 0$ then $g_1(y) = 0$, $g_2(y) \neq 0$.

So, assuming that the method is stable, it is of second order and Stetter's Theorem shows that

$$\zeta_n^N = y(x_n) + N^{-2}e(x_n) + O(N^{-3})$$

where $G'(y)e = -g_2(y)$

i.e. $e(a) = 0$

$$e'(x) - f(y(x))e(x) = -g_2(y(x)) \quad x \in [a, b].$$

To prove stability we assume that f satisfies a Lipschitz condition

$$\|f(y_1) - f(y_2)\| \leqslant L\|y_1 - y_2\|.$$

Suppose that, omitting suffix and superfix Ns

$$\Phi(\zeta) = d, \qquad \Phi(\xi) = e$$

and use the norms $\|\zeta\| = \max_{0 \leqslant n \leqslant N} \|\zeta_n\|$, $\|e\| = \max_{0 \leqslant n \leqslant N} \|e_n\|$, etc.

Then

$$||\zeta_0 - \xi_0|| = ||d_0 - e_0||$$

and

$$||\zeta_{n+1} - \xi_{n+1}|| = \left\| \begin{matrix} \zeta_n + \tfrac{1}{2}h[f(\zeta_n) + f(\zeta_n + hf(\zeta_n))] + d_n \\ -\xi_n - \tfrac{1}{2}h[f(\xi_n) + f(\xi_n + hf(\xi_n))] - e_n \end{matrix} \right\|$$
$$\leqslant (1 + hL + \tfrac{1}{2}h^2L^2)||\zeta_n - \xi_n|| + h||d_n - e_n||$$

for $0 \leqslant n \leqslant N$, using the Lipschitz condition.

It is now easily shown (cf. Henrici (1962 p. 18)) that:

$$||\zeta_n - \xi_n|| \leqslant \exp(nhL)||d_0 - e_0|| + (\exp(nhL) - 1).\max_{0 \leqslant i < n} ||d_n - e_n||/L.$$

Hence,

$$||\zeta - \xi|| \leqslant \max\{\exp((b-a)L), [\exp((b-a)L) - 1]/L\}.||d - e||$$

showing that (Φ_N^{-1}) is stable.

4. Deferred correction

The following theorem supplies some theoretical justification for the methods of deferred or difference correction (Fox and Goodwin (1949), Fox (1962)).

The problem considered is that of the previous paragraph and the initial discretization is as before, but a sequence of corrections of increasingly higher order is given by the $\psi_{N,i}$, $i = 1, 2, \ldots, q$.

Theorem Suppose that

(*a*) $Gy = b$.

(*b*) Φ_N is consistent of order p with G and that the sequence of inverse mappings (Φ_N^{-1}) is stable.

(*c*) $||\psi_{N,i}(\zeta_1^N) - \psi_{N,i}(\zeta_2^N)|| \leqslant KN^{-p}||\zeta_1^N - \zeta_2^N|| \quad 1 \leqslant i \leqslant q. \quad (1)$

(*d*) $||\Phi_N(\Delta_A^N Z) + \psi_{N,i}(\Delta_A^N Z) - \Delta_B^N G(Z)|| \leqslant CN^{-r_i} \quad 1 \leqslant i \leqslant q. \quad (2)$

i.e. that $\Phi_N + \psi_{N,i}$ is consistent of order r_i.

(*d*) Define η_i^N by

$$\left.\begin{aligned} &\Phi_N(\eta_0^N) = \Delta_B^N b \\ &\Phi_N(\eta_i^N) = \Delta_B^N b - \psi_{N,i}(\eta_{i-1}^N) \quad 1 \leqslant i \leqslant q. \end{aligned}\right\} (3)$$

Then
$$||\eta_i^N - \Delta_A^N y|| \leqslant C_i N^{-s_i} \quad (4)$$

where $s_0 = p$, $s_i = \min(r_i, s_{i-1} + p) \quad 1 \leqslant i \leqslant q$.

(Note that if $r_i = (i+1)p$ then also $s_i = (i+1)p$.)

PROOF. The consistency and stability show that there is a C_0 such that (4) holds for $i = 0$. Assume there are constants C_i such that (4) holds for all i such that $0 \leqslant i < j < q$.

Then

$$\begin{aligned} L||\eta_j^N - \Delta_A^N y|| &\leqslant ||\Phi_N(\eta_j^N) - \Phi_N(\Delta_A^N y)|| \quad \text{(by stability)} \\ &= ||\Delta_B^N b - \psi_{N,j}(\eta_{j-1}^N) - \Phi_N(\Delta_A^N y)|| \quad \text{(by (3))} \\ &\leqslant ||\psi_{N,j}(\eta_{j-1}^N) - \psi_{N,j}(\Delta_A^N y)|| \\ &\quad + ||\Delta_B^N G(y) - \Phi_N(\Delta_A^N y) - \psi_{N,j}(\Delta_A^N y)|| \\ &\leqslant K.N^{-p}C_{j-1}N^{-s_{j-1}} + CN^{-r_j} \end{aligned}$$

using (2) and (4) for $i = j - 1$.

So (4) holds with $i = j$ and hence for all i $(1 \leqslant i \leqslant q)$, proving the theorem.

The result of this section is similar to those of Pereyra (1966).

Example 2

Consider the two-point boundary value problem

$$y'' - f(x, y) = 0, \quad y(a) - \alpha = 0, \quad y(b) - \beta = 0.$$

This problem has been studied by Henrici (1962), Chapter 7, and by Lees (1965) and Pereyra (1966). We assume that $f(x, y)$ and $f_y(x, y)$ are continuous and $0 < f_y(x, y) \leqslant L$ for $a \leqslant x \leqslant b$ and all y, under which conditions the problem has a unique solution $y(x)$. (The existence of a unique solution can also be demonstrated under less stringent conditions than these (Lees (1965)).)

Take $A = C_R([a, b])$, $B = C_R([a, b]) \times R^2$, $A^N = B^N = R^{N+1}$, define Δ_A as before, take

$$\Delta_B^N\{Z_0, Z_1, Z_2\} = \{Z_0(x_1), Z_0(x_2), \ldots, Z_0(x_{N-1}), Z_1, Z_2\}.$$

and define Φ_N by

$$\Phi_N(\zeta^N) = \begin{cases} (\zeta_{n-1}^N - 2\zeta_n^N + \zeta_{n+1}^N)/h^2 - f(x_n, \zeta_n), & 0 < n < N \\ \zeta_0 - \alpha \\ \zeta_N - \beta. \end{cases}$$

Now Henrici (1962) gives a result (7–46 page 363), that implies immediately that, in our notation,

$$||(\Phi^{-1})'(\Delta_A^N Z)|| \leqslant \tfrac{1}{4}(b - a).$$

An almost identical analysis shows that

$$||\Phi^{-1}(b_1) - \Phi^{-1}(b_2)|| \leqslant \tfrac{1}{4}(b - a)||b_1 - b_2||,$$

and hence (Φ^{-1}) is stable.

The operators $\psi_{N,i}$ are defined by

$$\psi_{N,i}(\zeta^N) = \{\psi_{N,i,0}, \psi_{N,i,1}, \ldots, \psi_{N,i,N}\}$$

where

$$\psi_{N,i,0} = \psi_{N,i,N} = 0 \qquad i \geqslant 1$$

$$\psi_{N,1,n} = \frac{-1}{12}\delta^4\zeta_n^N/h^2. \qquad 2 \leqslant n \leqslant N - 2$$

$$\psi_{N,2,n} = \left[-\frac{1}{12}\delta^4\zeta_n^N + \frac{1}{90}\delta^6\zeta_n^N\right]/h^2,$$

etc., where terms are those of the expansion

$$h^2 w_r'' = \delta^2 w_r - \frac{1}{12}\delta^4 w_r + \frac{1}{90}\delta^6 w_r - \frac{1}{580}\delta^8 w_r + \ldots$$

and δ is the central difference operator given by

$$\delta w_r = w_{r+1/2} - w_{r-1/2}.$$

The above central difference expansions cannot be used near the ends of the range as values of ζ_n^N outside the range $0 \leqslant n \leqslant N$ are required. They must be replaced by unsymmetric formulae giving a similar accuracy, but using only the available ζ_n^N.

It is easily verified, using Taylor expansions that the theorem of (4) can be applied with $p = 2$ and $r_i = 2i + 2$ and hence if η_i is calculated according to the specification given there it will be correct to $O(N^{-2i-2})$ or $O(h^{2i+2})$.

References

BULIRSCH, R., and STOER, J. (1966). Numerical Treatment of Ordinary Differential Equations by Extrapolation Methods, *Num. Math.*, Vol. 8, pp. 1–13.

BULIRSCH, R., and STOER, J. (1966a). Asymptotic Upper and Lower Bounds for Results of Extrapolation Methods, *Num. Math.* Vol. 8, pp. 93–104.

BUTCHER, J. C. (1966). On the Convergence of Numerical Solutions to Ordinary Differential Equations, *Math. Comp.*, Vol. 20, No. 93, pp. 1–10.

COLLATZ, L. (1966). *Functional Analysis and Numerical Mathematics*, Academic Press, New York and London.

DIEUDONNÉ, J. (1960). *Foundations of Modern Analysis*, Academic Press, New York and London.

FOX, L. (1962). *Numerical Solution of Ordinary and Partial Differential Equations*, Pergamon, New York and London.

FOX, L., and GOODWIN, E. T. (1949). Some new methods for the numerical integration of ordinary differential equations, *Proc. Camb. Phil. Soc.*, Vol. 45, pp. 373–388.

GRAGG, W. G. (1965). On Extrapolation Algorithms for Ordinary Initial Value Problems, *J. SIAM. Numer. Anal. Ser. B*, Vol. 2, pp. 384–403.

HENRICI, P. (1962). *Discrete Variable Methods in Ordinary Differential Equations*, Wiley, New York and London.

LEES, M. (1966). Discrete Methods for Non-Linear Two-Point Boundary Value Problems, *Numerical Solution of Differential Equations* (ed. Bramble, J. H.), Academic Press, New York and London.

LIUSTERNIK, L. A., and SOBOLEV, V. J. (1961). *Elements of Functional Analysis*, Ungar, New York.

PEREYRA, V. (1966). On Improving an Approximate Solution of a Functional Equation by Deferred Corrections, *Num. Math.*, Vol. 8, pp. 376–391.

SIMMONS, G. F. (1963). *Introduction to Topology and Modern Analysis*, McGraw-Hill.

STETTER, H. J. (1965). Asymptotic Expansions for the Error of Discretization Algorithms for Non-linear Functional Equations, *Num. Math.*, Vol. 7, pp. 18–31.

STETTER, H. J. (1965a). A study of strong and weak stability in discretization algorithms, *J. SIAM. Numer. Anal. Ser. B*, Vol. 2, pp. 265–280.